
*University of Alberta*
*Printing Department*

THE UNIVERSITY OF ALBERTA

# A LINEAR PROGRAMMING APPROACH TO THE INTERPRETATION OF EARTH RESISTIVITY DATA

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

DEPARTMENT OF PHYSICS

by

Syed Amir Bukhari

Edmonton, Alberta

May, 1961

## Acknowledgements

I am grateful to Dr. Vozoff, who suggested and supervised this work. I am deeply indebted to him for his many valuable suggestions and continued interest in the work.

This work was supported in part by The Geological Survey of Canada, Department of Mines and Technical Surveys. It is with much pleasure that I acknowledge the same.

The members of the staff of the Computing Center at the University of Alberta provided valuable assistance and I wish to express much appreciation for the same.

Miss G. Tratt typed the manuscript and Mr. R. Hunt prepared the drawings. I wish to thank them for their excellent work.

Table of Contents

## Abstract

The problem of direct interpretation of earth resistivity data was attacked by Ness (4) using a method of least squares. The same problem is considered here, using linear programming methods for its solution. These methods were tried because they allow data fitting in other than a least square sense.

Two sets of model data were analyzed, and the results are comparable with those obtained by Ness, both in quality of fit and in computation time required. It appears that both methods reduce errors to within the range of the accuracy of the measurements and of the approximations.

The linear programming methods possess a decided advantage in their flexibility, which allow subjective control where desired, and which would also allow the use of information from other sources with little modification of the programs.

# I. Introduction

The direct interpretation of resistivity data requires the solution of an inverse boundary value problem. The main work in this connection was done by Vozoff [6,7] and Ness [4]. In [4] one finds a considerable historical account of resistivity measurements, and of attempts at direct and indirect interpretation.

The direct interpretation problem has been formulated by a first approximation, as a set of linear equations and adapted for computer solution [4] [6,7]. The method of least squares was used in fitting the observed data. It was indicated that this method gave results which may be classed as semiquantitative. Detailed discussion on this aspect of the problem is found in [4,6,7].

Linear programming is, comparatively, a new mathematical method for solving problems concerned with decision-making. The method of solution is basically iterative, and thus ideally suited for electronic computers. Further, the flexibility of its methods give a powerful control over the solution itself. A large variety of problems are amenable to analysis by linear programming. It can be used, for example, to minimize the sum of absolute value of errors, or to emphasize or ignore the effect of a number of variables.

In the following, an attempt is made to use the method of linear programming in interpreting the resistivity data directly.

## II. Formulation of the Problem

The problem of interpreting the apparent resistivity data, assuming general inhomogeneity is fundamentally based on the following hypothesis, originally put forward by Vozoff [6].

"The region of interest can be considered as being made up of small homogeneous blocks of given geometry but unknown conductivity."

This leads to a three or two dimensional array of regular blocks of known geometry. The simplest shape to choose would be that of a cube. Accordingly in the following we shall assume the subsurface region of interest to be made up of an array of small homogeneous cubes having unknown conductivities $\sigma_i$, $i = 1, 2, \ldots n$, where n is the total number of cubes into which the subsurface region of interest has been divided. This matrix of n cubes is assumed to be embedded in an otherwise semi-infinite homogeneous medium characterized by $\sigma_0$, the background conductivity.

Suppose $\Omega$ is an apparent resistivity reading, at a particular location on the surface of the earth, or on the model in the laboratory. We assume the combined effect of the n cubes together with the primary effect to be equal to $\Omega$.

Thus symbolically we have,

$$\Omega = \frac{1}{\sigma_0} + F(\sigma_0, \sigma_1, \ldots \sigma_n, x, y, z)$$

where F is some general function of the arguments $\sigma_i$, $i = 1, 2, \ldots n$, and x, y, z.

The approximation is such that it linearizes the system and expresses F as

$$F(\sigma_0, \sigma_1 \ldots \sigma_n, x, y, z) \simeq \sum_{i=1}^{n} G_i(x, y, z) X_i(\sigma_i, \sigma_0)$$

where

$X_i(\sigma_i, \sigma_0)$ = function of $\sigma_i$ and $\sigma_0$ only

$G_i(x, y, z)$ = geometric factor associated with $i^{th}$ cube and the relative positions of source and receiver.

We shall state here the final results of the first approximation, referring for actual details to the references cited.

The linearized system of equations for pole-pole configuration (see Fig. 1) i.e. for the potential at a point P on the surface due to a unit current source at another point Q on the surface is

$$[\varphi_m] \simeq [\varphi_m^0] + \frac{c}{4\pi^2\sigma_0} [G_{mi}][X_i] \qquad (1)$$

GEOMETRICAL PICTURE FOR A SINGLE CUBE

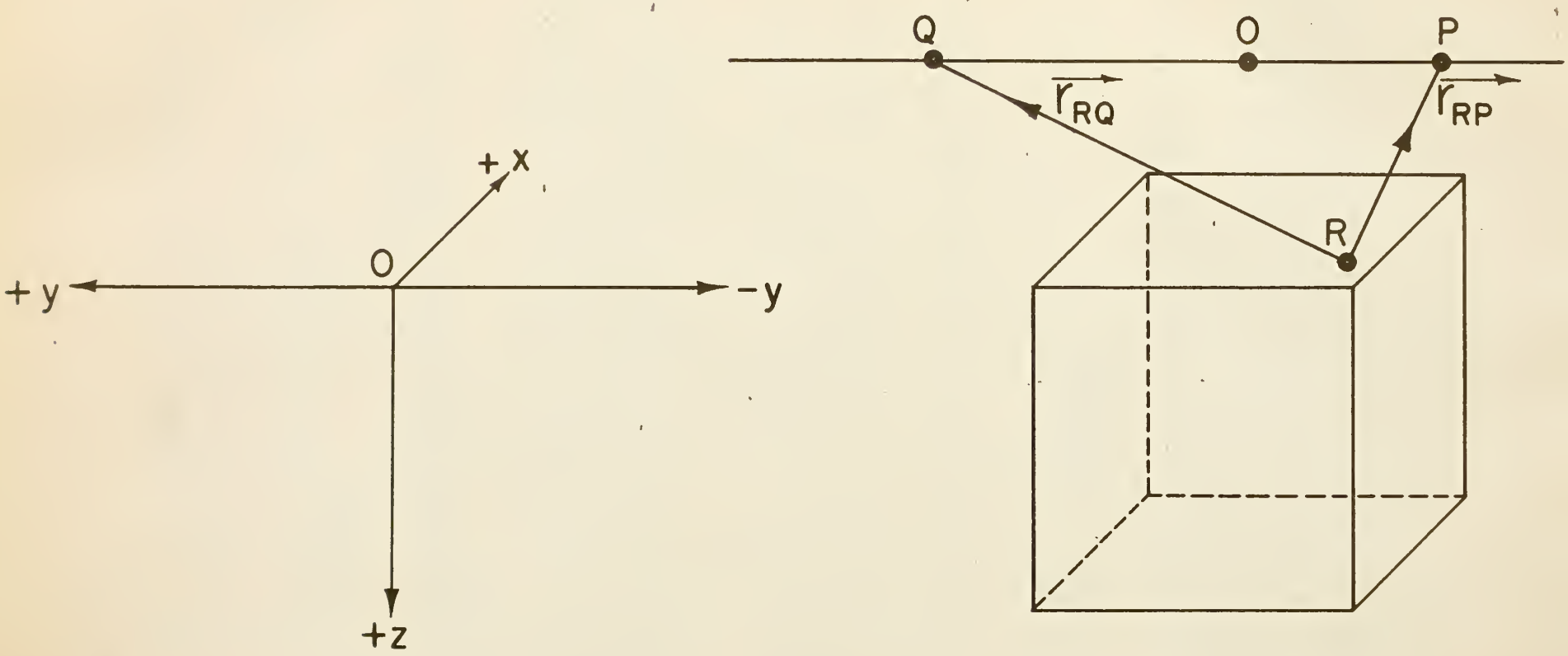



FIG. 1

We shall use

(a) index i to refer to the cubes into which the subsurface region has been divided. (i = 1, 2, ....n)

(b) index j to refer to receiver point P. (In numerical calculation values j = -5, -4, -3, .....3, 4, 5 were given).

(c) index k to refer to source point Q. (In numerical calculations values k = -5, -4, .....4, 5 were given).

An apparent resistivity reading is characterized by a combination of j and k index. We shall denote the combination jk by m. Thus m refers to apparent resistivity readings and whenever convenient we shall replace m by $(jk)$ and vice versa, e.g.

$$G_{mi} \equiv G_{ijk}$$

In (1): $[\varphi_m]$ is an m x 1 matrix. Matrix elements $\varphi_m$ represent the total potential for a combination of j and k (which is denoted by m).

$[\varphi^\circ_m]$ is an m x 1 matrix. Matrix element $\varphi^\circ_m$ is given by

$$\varphi^\circ_m = \frac{c}{2\pi\sigma_0 r_{jk}}$$

and may be interpreted as the primary potential due to a unit source at k and receiver at j position.

$[\ G_{mi}\ ]$ is an m x n matrix. $G_{mi}$ is the approximated geometric factor associated with cube i and source-receiver

combination designated by m. For its explicit form refer to Appendix I.

$[X_i]$ is an $n \times 1$ matrix.

$$c X_i = c \frac{\sigma_i - \sigma_0}{\sigma_i + 2\sigma_0}; \quad c = 3{\cdot}6 \qquad (2)$$

represents the conductivity contrast factor for $i^{th}$ cube.

Approximation (1) assumes that

(a) There is a very thin layer of uniform conductivity $\sigma_0$ on the surface of the earth. Thus $\sigma_j = \sigma_k = \sigma_0$ a constant, where $\sigma_j$ = conductivity at receiver position j; $\sigma_k$ = conductivity at source position k.

(b) The electric field at a conductivity interface is just the primary electric field.

(c) The form of $X_i$ as given by (2) does not follow from the approximation itself. It was suggested by Vozoff, and $c = 3.6$ was empirically determined by him [7]. It may be mentioned that the actual $X_i$ which appears in the approximation is given by

$$X_i = \ln \frac{\sigma_i}{\sigma_0}$$

However, the behaviour of $\ln \frac{\sigma_i}{\sigma_0}$ as a conductivity contrast factor is contrary to experience. It appears that $X_i$ may be an algebraic function. Hence the empirical suggestion by Vozoff [7].

The interpretation scheme (1) is based on the fundamental hypothesis mentioned earlier. The hypothesis is moderately

successful in actual practice. The reason for this may be seen by examining the physical situation. Suppose we have only one disturbing cube. According to the hypothesis we have to define a region of interest, and divide it into a regular array of cubes. In the solution of (1), the disturbing cube shows up as such, and the remaining cubes show up as either having conductivity equal to $\sigma_0$ or very slightly different from it. This behaviour makes it possible to get a reasonable fit. Since (1) is not exact, a situation somewhat like the above is really intuitively suggested. If (1) was exact and the physical situation perfect, then and only then would we expect the disturbing cube to be the only one to show up in the solution.

In applying this hypothesis, several practical problems must be considered. One concerns the optimal size of the blocks into which the subsurface is to be subdivided. This should be governed by electrode spacing, and should vary with distance (horizontal as well as vertical) from the electrode traverse. Another concerns the amount of information available from a given number of readings. Thus it may be asked, that for a given number m of readings on the surface of the earth, what should be the number n of subsurface blocks of which $k \leq n$ are anomalous, so that all the anomalous blocks may be completely determined and vice versa. These and similar considerations are very important, and obviously will

have a great bearing on devising optimum prospecting plans.

Vozoff [6,7] and Ness [4], have discussed some of these questions in detail.

It is interesting, however, to look at the last question from the point of view of information theory.

Suppose there are n identical coins and exactly one of these is false, being either lighter or heavier than the good coins. It is required to detect this coin in a minimum number of weighings m, by an equal arm balance.

In reference [8] there appears a solution to this problem. The solution is based on the following assumptions:

1. The permutations of the states of n coins are distinguishable. This can be done by labelling the coins as 1, 2, ...., n. (This may appear to conflict with the word "identical" in the problem, but it need not bother us since a labelling scheme will be used in the application).
2. Any one of the coins could be false.

The following solution is given:

$$n \leq \frac{3^m - 3}{2}$$

Thus to detect a false coin from a set of 12 coins, 3 or more weighings are necessary.

Assuming any $k \leq n - 1$ coins are false, the above formula is extended to take the form

$$n \leq \frac{3^{\frac{m}{k}} - 3}{2} \qquad (3)$$

Suppose now we have a group of n cubes embedded in a uniform medium of conductivity $\sigma_0$ , and that the conductivities of k of these cubes differ from $\sigma_0$ . The conductivities are designated by $\sigma_i$ , $i = 0, 1, 2, \ldots, n$, and we can use equation (3) to relate n, the number of cubes, m the number of observations and k the number of "anomalous" cubes. In the resistivity problem usually k = n, and in all we have n + 1 conductivities $\sigma_i$ , $i = 0, 1, \ldots n$. Hence in (3) we put k = n, and replace n by n + 1 to get

$$n \leq \frac{3^{\frac{m}{n}} - 5}{2} \qquad (3a)$$

(where n and m are integers) and assume that it gives an indication of the minimum number of readings m, required to determine the conductivities of n regions.

In typical situations where perhaps 30 to 50 observations have been taken along a traverse, one should not expect to be able to resolve more than about one-third of this number of cubes.

## III. Errors and Their Estimates

A measured number by itself signifies nothing that could be safely interpreted theoretically. If an aggregate is available, only then can the theoretical significance of the measurements be assessed. For in that case only have we the means for determining the error, or measure of precision, of results.

A "true" value of a measured quantity cannot be determined. What we accept empirically as a final value is really a maximum likelihood value. The concept of maximum likelihood attains meaning only if a sufficient number of measurements is available. From this it is possible to plot a distribution curve that represents the information contained therein. The procedure more commonly followed is to assume a Gaussian distribution and to apply least squares fitting to determine a most likely value. The main reason for the wide usage of the least squares, method is its simplicity and the ease with which it can be applied. However, it must be pointed out explicitly that one cannot always apply the method of least squares, and for certain types of observations the method will not determine a most probable value, and indeed may determine a least probable value. An example illustrating this point may be found in [5] on pages 217 - 218.

In resistivity work magnitude determination involves reading the coincidence of a pointer on a scale. The limitations of an instrument and our senses necessarily imply an

inherent uncertainty in the determined magnitude. These errors may be termed the resultant experimental error. These errors involve, really, a rounding off of the magnitude, and are also termed roundoff errors.

The observed data are usually approximated by theoretical expressions based on certain hypotheses. This introduces, what we may call theoretical or truncation error. A truncation error corresponds to the fact that, whereas an exact result would be afforded by an infinite sequence of steps, the process is "truncated" after a certain finite number of steps. (Indeed the linear approximation given by equations (1) is got by truncating the process of successive approximations, after a first approximation has been made.)

In using a theoretical approximation to extract useful information from observed data, the question of estimating errors becomes important. The useful information which a theoretical approximation is designed to extract usually consists of numerical values of certain parameters. In particular these are the numerical values of $X_i$ for $i = 1, 2, \ldots, n$ in our case. The question arises, if a number of determinations have given

$$a_1, a_2, \ldots, a_m$$

as values of a certain quantity in various situations, what can we say (in a maximum likelihood sense) about the parameters

which appear in the theoretical approximation of this quantity? In particular if $a_1, a_2, \cdots a_m$, were the observed potentials $\varphi_1, \varphi_2, \cdots \varphi_m$ then what can we say about $X_i$ ?

It turns out that if $a_1, a_2, \cdots a_m$ are given without any indication of the way in which they were arrived at, a definite estimate in a maximum likelihood sense can only be made under certain arbitrary assumptions. From these assumptions it is possible to find a law of distribution of errors.

If the law of distribution of error is known or assumed, and the resulting fit is not good, then necessarily we must re-examine the theoretical approximation. It is however usual that the law of error is unknown, and one has to assume its form.

In practice, one assumes the law of error (usually Gaussian Law of error) then develops a theoretical approximation, so as to arrive at a good fit. Clearly this is an apriori procedure, and depends largely in one's belief in the assumed law of error.

In the following some error estimates, the underlying assumptions, and the corresponding laws of error are stated. Also we shall mean by "error" the difference between observed and predicted values of the quantity measured.*

1. If we assume that for a given set of data, the best possible estimate of the measured quantity is given by minimizing the sum of the squares of the errors, then this assumption implies that the probability density of the errors

*The more common term for this difference, in statistical analysis, is the "residual."

is given by an expression of the form

$$y = \frac{h}{\sqrt{\pi}} e^{-h^2x^2}$$

where y is the probability density of the distribution, and is such that the probability of an error in the determination lying between x and x + dx is equal to y dx. h is a constant called the precision constant. Fisher [2] has shown that if the above assumption is strictly obeyed then minimizing the sum of the squares of the errors gives maximum information possible, for a given number of readings. He has further shown that if the above law is valid, but the sum of the absolute value of errors is minimized, then approximately 14% more readings are required to obtain the same accuracy which may be obtained by minimizing the sum of the squares of the errors.

Burnside [1] has shown that the above assumption is reasonable under the following conditions:

(i) If actual error is made up of a sufficiently large number of minor errors of the same magnitude, each of which is equally likely to be positive or negative, then it follows Gauss's Law.

(ii) Error, if it arises as the sum of a sufficiently large number of minor errors of which the probable value is zero, follows Gauss's Law.

(iii) If a number of independent component errors all follow Gauss's Law, and if the resultant error is compounded linearly from them in any way, then it also follows Gauss's Law.

It is possible, however, that each component of resultant error may not follow Gauss's Law. Though this law is very popularly used, yet one must exercise care in its use.

2. Assume that minimizing the absolute sum of errors gives the best possible fit.

Suppose n measurements are made to estimate a certain magnitude whose true value is $\gamma$. Let these n values be $a_1, a_2 \ldots a_n$ arranged in ascending order of magnitude. Assume further that $\gamma$ lies between $a_r$ and $a_{r+1}$. The sum of absolute values of error is

$$(2r-n)\gamma + \sum_{K=r+1}^{n} a_K - \sum_{K=1}^{r} a_K = \epsilon$$

| | | |
|---|---|---|
| when | $2r > n$ | $\epsilon$ is least for $\gamma = a_r$ |
| | $2r < n$ | $\epsilon$ is least for $\gamma = a_{r+1}$ |
| | $2r = n$ | $\epsilon$ is independent of $\gamma$ |

When n is odd, say equal to $2m' + 1$, $\epsilon$ is least when $\gamma = a_{m'}$. When n is even, say $2m'$, $\epsilon$ is constant for all values of $\gamma$ from $a_{m'}$ to $a_{m'+1}$ and this value is least.

If the above assumption is justified then the law of error is of the form

$$y = \frac{h}{2} e^{-h|x|}$$

and it can be shown (Fisher: 2) that the above estimate is such that for a given number of readings, maximum information is obtained by minimizing the absolute sum of errors.

Suppose errors in a system of equations are denoted by $\epsilon_i$, index $i$ referring to $i$th equation. $\epsilon_i$ can be either positive, zero, or negative. We can express $\epsilon_i$ as

$$\epsilon_i = \alpha_i - \beta_i$$

where all $\alpha_i$ and $\beta_i$ are either positive numbers or zero, and only one of the pair $\alpha_i, \beta_i$ can be non-zero.

The above shows that when dealing with errors as such, we can have an estimator of error, viz:

$$\sum_i |\epsilon_i| = \sum_i (\alpha_i + \beta_i) \equiv |\epsilon|$$

Minimizing $|\epsilon|$ means that the sum of absolute value of errors is minimized.

3. Another estimate may be that the best value is got by minimizing the $2m^{th}$ power of errors [1].

$$\sum_{i=1}^{n} (a_i - \gamma)^{2m-1} = 0$$

It can be shown [1], that when $m > 1$

$$\gamma = \frac{1}{2}(a_1 + a_n)$$

where $a_1 = \underset{i}{\text{minimum}}\ a_i$

$$a_n = \underset{i}{\text{maximum}}\ a_i$$

This shows that this estimate makes the determination of true value depend more on larger and smaller values observed.

Each datum of apparent resistivity is got by reading voltmeters and ammeters, used in the instrumentation. Assuming these meters to be properly calibrated, we may expect the errors resulting from the use of these meters to be Gaussian. In the calculation of the geometrical coefficients $G_{mi}$ an error of the order of 1% is indicated. Whether or not this is Gaussian is open to question. Further it is also not quite clear if the error introduced by the first approximation used to arrive at equations (1) is Gaussian. This means that the least squares fitting of the data may or may not be the best approach. In view of the semiquantitative results got by using the method of least squares, in the following an attempt is made to use a fitting approach in which the sum of the absolute values of the errors is minimized. Before proceeding any further an example is given, showing that minimizing the sum of the squares of the errors may not always minimize the total error in the solution. The total error is defined as the sum of the absolute value of the errors.

In the example the data $f(x)$ were required to be fitted to a straight line. The two solutions are shown in Table I.

TABLE I

| X | 0 | 1 | 2 | 3 | 4 | Total absolute error |
|---|---|---|---|---|---|---|
| Given $f(x)$ | 1.00 | 3.85 | 6.50 | 9.35 | 12.05 | |
| L.S. | 1.03 | 3.79 | 6.55 | 9.31 | 12.07 | .20 |
| L.P. | 1.00 | 3.76 | 6.53 | 9.29 | 12.05 | .18 |

L.S. ≡ Least squares approximation to $f(x)$

L.P. ≡ Linear Programming approximation to $f(x)$ where sum of the absolute value of the errors was minimized.

## IV. Linear Programming

Linear Programming describes the solution of the following problem:

Find a column vector $X = (X_1, X_2, \ldots X_n)$, which minimizes

$$f = cX$$

Subject to the conditions

$$\begin{aligned} GX &\geqslant b \\ X &\geqslant 0 \end{aligned} \qquad \text{I}$$

where

| | |
|---|---|
| $G = (G_{mn})$ | is m x n matrix |
| $b = (b_1 \ldots bm)$ | is a column vector |
| $c = (c_1 \ldots c_n)$ | is a row vector |

This is termed the 'primal' problem.

A solution satisfying I is termed a feasible solution. A basic feasible solution (b.f.s.) is a feasible solution with no more than m positive $X_i$. A b.f.s. which maximizes f is termed the optimal solution.

### Dual Problem

Associated with every primal problem is the 'dual' problem:

Find a row vector $W = (W_1 \ldots W_m)$ which maximizes

$$g = Wb$$

subject to the conditions

$$WG \leq c$$
$$W \geq 0 \qquad \text{II}$$

Solution of the primal problem provides useful information about the solution of the dual problem. In fact if $X^o$ is an optimal solution to I and $W^o$ optimal solution to II, then

$$\max f = cX^o = \min g = W^o b.$$

Standard methods exist for the solution of L.P. problems [9]. Perhaps most well known is the Simplex Method. However, the Revised Simplex Method [3] is more suitable for computer calculations and will be used for the numerical computations in this work.

The success of L.P. method, especially for large problems, depends on the use of fast electronic computers. The size of an L.P. problem which can be solved on an electronic computer is usually limited by the number of rows in G matrix. A problem, when formulated as a primal problem, may be too large for the computer. However, its dual problem may be of acceptable dimensions. As was mentioned earlier, one can extract the optimal solution of the primal problem, if the final iteration in the solution of the dual problem is known. It is to stress this point that the two problems are stated together.

The method of solution of L.P. problems is iterative in nature, and is easily adaptable for computer programming and

solution. This allows one to follow through the successive stages of the solution.

## V. Formulation of the Resistivity Interpretation Problem as a Linear Program

The resistivity approximation, for unit current, may be written as follows:

$$[\varphi_m] - [\varphi_m^\circ] \simeq \frac{c}{4\pi^2\sigma_o} [G_{mi}][X_i]$$

Putting

$$\frac{4\pi^2\sigma_o}{c} [\varphi_m - \varphi_m^\circ] = [b_m]$$

we can write

$$[G_{mi}][X_i] \simeq [b_m]$$

where

$$X_i = \frac{\sigma_i - \sigma_o}{\sigma_i + 2\sigma_o} \tag{3}$$

We introduce in the approximation a column vector $[\epsilon_m]$ and interpret it as an error column vector. Thus the uncertainty in $\simeq$ is removed, and we can write a system of equalities, instead of the approximate equations.

$$[G_{mi}][X_i] + [\epsilon_m] = [b_m]$$

$[\epsilon_m]$ has m elements.

The elements of $[X_1]$ and $[\epsilon_m]$ can be either negative or positive, while a linear program and its method of solution requires that we deal with positive variables only. We can, however, reduce the above system of equations to the canonical form of linear program constraints by expressing the elements of $[X_i]$ and $[\epsilon_m]$ as a difference of two positive numbers, thus

$$[X_i] = [X_i^+ - X_i^-], \quad [\epsilon_m] = [\alpha_m - \beta_m]$$

Finally we could write

$$[G_{mi}][X_i^+ - X_i^-] + [\alpha_m] - [\beta_m] = [b_m]$$

$$[X_i^+] \geq 0 \qquad [X_i^-] \geq 0$$

$$[\alpha_m] \geq 0 \qquad [\beta_m] \geq 0$$

and minimize

$$[C_m][\alpha_m + \beta_m] \equiv f' \qquad \text{III}$$

where $[C_m]$ is a row vector, all the elements of which are unity. $[C_m]$ has m elements.

III is a standard linear program.

In III it is required to minimize:

$$f' = [C_m][\alpha_m + \beta_m] = \sum_m C_m(\alpha_m + \beta_m) = \sum_m (\alpha_m + \beta_m)$$

which is the sum of absolute errors. In doing this we do not

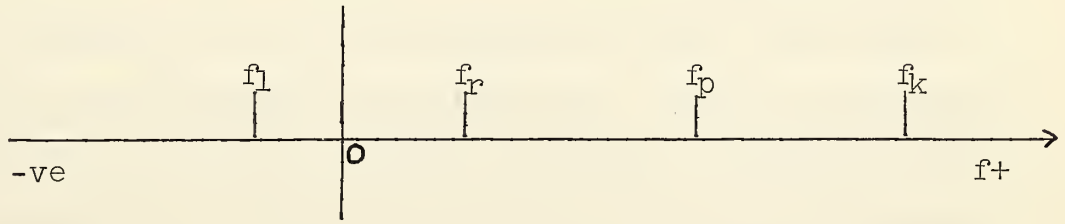
have any way to ensure that the sum of total positive error is the same as the sum of total negative error. However, we may want to design our procedure to equalize these errors. This can be done, using L.P. methods, by reducing the absolute numerical value of

$$[C_m][\alpha_m - \beta_m] \equiv f \qquad \text{(A)}$$

The simplex method first minimizes $[C_m][\alpha_m]$ (if not, it can easily be made to do so) say to a value $f_p$. Then it starts maximizing $[C_m][\beta_m]$, thus reducing $[C_m][\alpha_m - \beta_m]$ as a whole. It is quite possible that f may change sign. The method of solution is iterative, and we can easily have the computer print the solutions of each iteration. Each value of f printed will be smaller than the previous one. Suppose the sequence of the values of f is as

$$f_k \ldots\ldots f_p \ldots\ldots f_r f_l$$

where $f_r$ is the value of f just before f changes sign and $f_l$ its value immediately after it has changed sign. The sequence shown is in descending order of magnitude. Geometrically, we can show the above sequence as



We stop as soon as f has changed sign. Then optimum f is given by

$$\min \left\{ f_r, f_l \right\}$$

This, however, does not mean that the sum of absolute value of errors, which is given by $[C_m][\alpha_m + \beta_m]$ is also minimized. It may or may not have been minimized. By reducing the numerical value of $[C_m][\alpha_m - \beta_m]$ we have only ensured that the sum of errors entering the solution with a positive sign (i.e. sum of the non-zero elements of $[\alpha_m]$) is as closely as possible equal to the sum of errors entering the solution with a negative sign (i.e. sum of the non-zero elements of $[\beta_m]$).

Flexibility of the methods of L.P. is easily seen. Firstly, in the case when

$$f \equiv [C_m][\alpha_m - \beta_m]$$

there is always the possibility of properly combining the solutions corresponding to $f_r$ and $f_l$, which may improve the final solution considerably.

In cases when the limitations of the computer do not allow a complete run of the program, we still have quite a few ways to get around this difficulty. Firstly, perhaps a dual formulation of the problem may help. Secondly, the dual formulation failing, we may run the program in steps.

For convenience of writing denote $[X_i^+ - X_i^-]$ by $[X_i]$, keeping its significance in mind. Thus we may first run:

$$[G_{mi}][X_i] - [\beta_m] = [b_m]$$

and then run

$$[G_{mi}][X_i] + [\alpha_m] = [b_m]$$

Suppose solutions are $[X_{\beta i}]$ and $[X_{\alpha i}]$, then

$$[X_i] = \frac{[X_{\alpha i}] + \lambda [X_{\beta i}]}{1 + \lambda}$$

is a solution to the complete problem. By making a suitable choice of $\lambda$ it is possible to reduce the sum of absolute value of errors even more.

Fitting of data by successive runs is quite successful, and by devising a suitable procedure one can also exercise control to some extent, on the solution of the problem. This is quite important, especially when one desires to judge the physical situation subjectively and control the fitting of data accordingly. This is possible, because it is a feature of L.P. that it can be used to minimize or maximize the effect of any variable or variables.

Minimizing $[C_m][\alpha_m + \beta_m] = \sum_m (\alpha_m + \beta_m)$ does not necessarily mean individual errors $\alpha_k$ and $\beta_k$ are also being minimized. In the L.P. process of reducing the sum it is possible that a few of the errors actually become larger. However, it is also desirable that individual errors

must be made as small as possible. To achieve this it is necessary to fit the data in a systematically controlled way. For example, it was found that a problem solved by linearly combining the solutions of the two problems

$$[G_{mi}][X_i] - [\beta_m] = [b_m]$$

$$[G_{mi}][X_i] + [\alpha_m] = [b_m]$$

did not give an acceptable fit to our data. The reason for discarding the solution was that though the sum of the absolute value of error was in the range which could be accepted, some of the individual errors were quite large. The actual situation was that, of the twenty-two error variables, three or four variables accounted for 50 per cent of the total possible error. Accordingly some control was exercised to get a "smoother" fit. This was possible mainly because the fitting was done in successive stages, due to limitations of the computer.

The point, however, to emphasize is that it is quite possible to use L.P. methods effectively. Flexibility of L.P. methods are a decided advantage when it is desired to incorporate some subjective judgement into the solution.

In formulating the problem as given by III we have introduced (2n + 2m) variables, and we have only m equations,

and (2n + 2m) inequalities. The variables have been introduced as follows:

(i) m variables in $[\alpha_m]$. These are those error variables which enter the solution with a positive sign.

(ii) m variables in $[\beta_m]$. These are those error variables which enter the solution with a negative sign.

(iii) 2n variables in $[X_i^+ - X_i^-]$

In any solution, however, m elements of $[\alpha_m]$ and $[\beta_m]$ will be zero, since of $\alpha_k$ and $\beta_k$ only one can have a non-zero value, the other will be zero. The same applies to $X_k^+$, $X_k^-$, only one of the two corresponding elements will appear in the solution, the other will be zero. Thus the effective solution is reduced to choosing m variables out of n + m variables. Since n will be greater than one, the number of error variables in the solution will be m - n. If the solution progresses, so that all the n variables contribute to the solution then presumably the total error will be the least. However, as also mentioned in the end of Section II, it will be necessary to limit the maximum number of cubes, and hence the number n. This limitation as shown earlier may be of the order given by the relation

$$3n \leq m$$

The introduction of so many variables is largely due to the desire of solving the problem with minimum of subjective manipulation, especially where signs of the variables are concerned. However, in the case where n is not very large, a close examination of the problem will usually indicate correctly, which of $X_i^+$ or $X_i^-$ will enter the solution. In other words, when the number of cubes into which the subsurface region has been divided is not very large, then it is possible to decide by inspection of the equation whether a cube is resistive or conductive relative to the background conductivity. This assumes, of course, that the signs of the G coefficients are correct. This procedure reduces the number of variables and contributes to a great saving of time. For example, in the case of a single cube,

$$G_{il} X_l + \alpha_i - \beta_i = b_i$$

If the sign of $G_{il}$ is correct, then it is a physical necessity the sign of $G_{il} X_l$ be the same as $b_i$. This fixes the sign of $X_l$ uniquely, and one of $(X_l^+, X_l^-)$ can be discarded.

## VI. Model Data Analysis

Model data due to single disturbing cubes are analysed. We denote the resistivity contrast factor of $i^{th}$ cube by $X_i (i = 1, \ldots n)$. (Since we will not have occasion to use the X axis explicitly, there need be no confusion.)

The model experiments were carried out in a wooden tank by Dr. Vozoff. The conductivity of saline water was properly adjusted and a resistive or conductive cube lowered on a stand in the tank. The cube was of 2" side. Pole-dipole measurements were carried out with the dipole spacing being 2". [7]

In the actual L.P. analysis of observed data, the disturbing cube $X_1$ was "surrounded" by five other cubes shown in Fig. 2 and distinguished as $X_2$, $X_3$, $X_4$, $X_5$, $X_6$.

The standard graphical presentation of data is shown in Fig. 3 [7] Data were available for all the points marked ⊗ as well as the points numbered from 1 to 22. The data referring to positions ⊗ were not used mainly because the coefficients $G_{ijk}$ were not available. The data at the other 22 points marked as 1, 2, .... 22 were ultilized. To each of these datum points there corresponds a constraint equation. We shall use index m = 1, 2, .... 22 to denote these points and the corresponding equations. Thus in the L.P. Problem we have

MATRIX OF CUBES

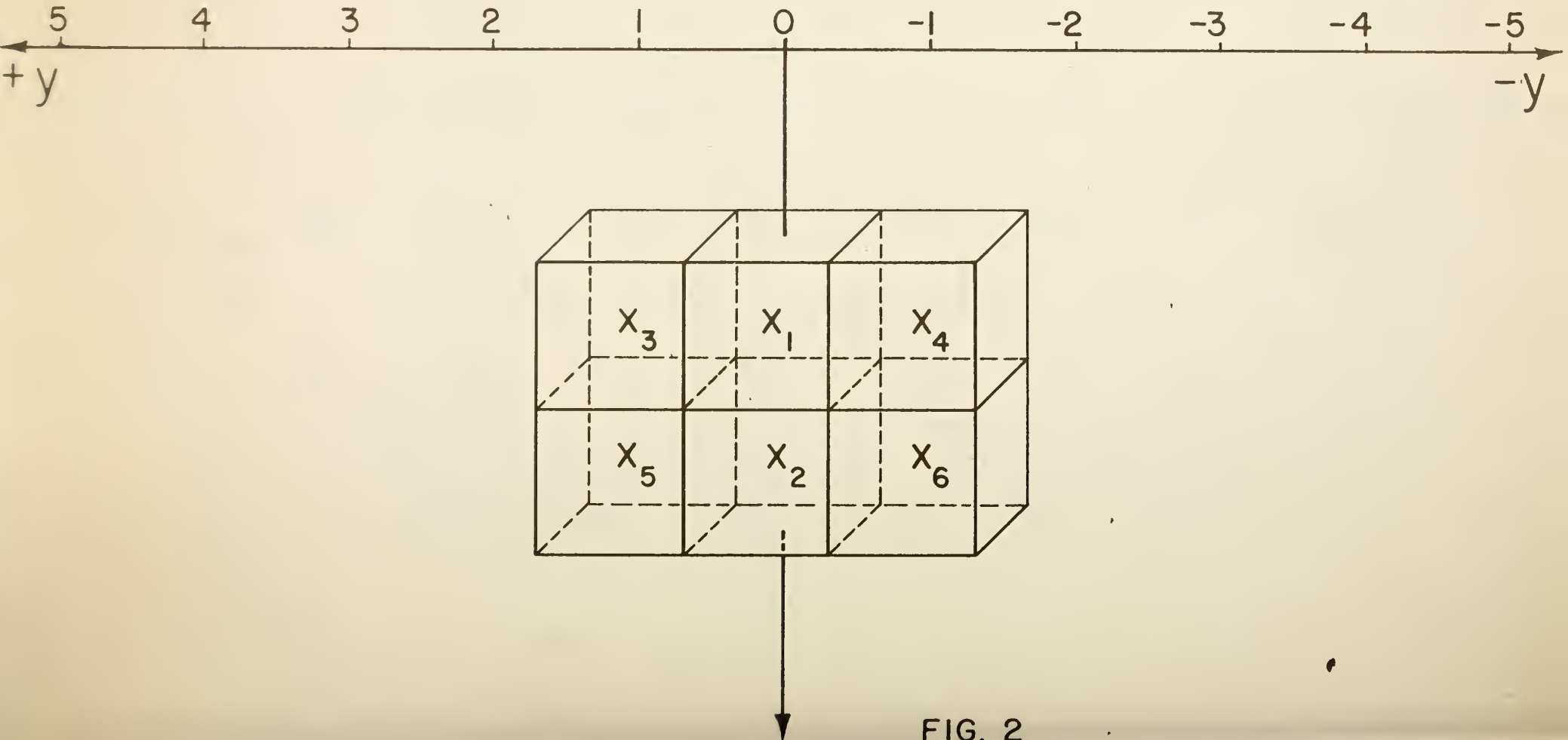



FIG. 2

22 equations. The combination of J and k, and hence the m to which these equations refer are tabulated with the results of the analyses of the data. Only J is given, it being assumed that the other end of the dipole is on location (j - 1). A typical pole-dipole configuration is shown in Fig. 3, where the receiver is at (2,1) and source at (4). This combination of source and receiver positions is assigned an index m = 1. Thus this combination in the tables I - VI appears as j = 2, k = 4, m = 1. In the figures only the m index is shown, under the black dot.

Resistive Cube

Following are the results of analysis on the data when the model tank had a single resistive cube as the disturbing element.

Due to the ultimate size of L.P. problem, it was decided to successively approach the solution.

From the nature of approximation being used it is evident that the cubes to which $X_1$, $X_3$, $X_4$, refer are the most important. Hence in the first step a solution is approximated using only these cubes and the L.P. equations:

$$[G_{mi}][X_i] - [\beta_m] = [b_m] \qquad \text{Step I.}$$

Here we are forcing all errors to be negative, and are denoting them by $[\beta_m]$. Further along in the solution, when both positive and negative errors are allowed, they will be denoted by $[\epsilon_m]$.

The numerical results of this step are given in Table II. It will be noted that all errors are entering as $-\beta_m$. This effectively means that $X_1$ has been assigned its upper limit. For a better fit $X_1$ must assume a lower value. This, however, will require that some errors enter the solution with positive sign. This may be seen as follows:

Suppose $X_1 = X_1' + X_1''$ $\qquad X_1' > X_1''$

The $i$th equation is

$$G_{i1}X_1 + G_{i3}X_3 + G_{i4}X_4 - \beta_i = b_i$$

Therefore $$G_{i1}X_1' + G_{i3}X_3 + G_{i4}X_4 + (G_{i1}X_1'' - \beta_i) = b_i$$

When $X_1 \longrightarrow X_1'$ in a succeeding iteration,

then $$-\beta_i \longrightarrow (G_{i1}X_1'' - \beta_i) = \epsilon_i$$

$\epsilon_i$ can be positive or negative.

It is easily seen that numerical value of errors can be decreased by properly choosing $X_1''$. Two different approaches were used to choose $X_1''$. These are labelled as A and B. Approach A was used in the analysis of resistive cube data now under consideration. Approach B was used to analyse the data from a conductive cube. However, for the sake of continuity both of these approaches are described below.

A. In Step I of the resistive cube analysis $\beta_9, \beta_{22}$ are zero. It is easy to find an upper value of $X_1''$ such that when $X_1 \longrightarrow X_1'$, errors corresponding to

TABLE II

Partial Solution, Resistive Cube

| k | j | m | $b_m$ | $\beta_m$ |
|---|---|---|---|---|
| 4 | 2 | 1 | 5.56 | 1.20 |
| 4 | 1 | 2 | 3.33 | 3.87 |
| 4 | 0 | 3 | 7.50 | 6.97 |
| 4 | -1 | 4 | 0.555 | 0.220 |
| 3 | 1 | 5 | 2.78 | 6.53 |
| 3 | 0 | 6 | 14.4 | 10.5 |
| 3 | -1 | 7 | 2.22 | 0.80 |
| 3 | -2 | 8 | 1.39 | 0.77 |
| 2 | 1 | 9 | 8.33 | 0.00 |
| 2 | 0 | 10 | 27.8 | 21.9 |
| 2 | -1 | 11 | 7.22 | 2.36 |
| 2 | -2 | 12 | 0.555 | 1.05 |
| 2 | -3 | 13 | 1.11 | 1.04 |
| 1 | 0 | 14 | 75.0 | 36.1 |
| 1 | -1 | 15 | 25.0 | 12.1 |
| 1 | -2 | 16 | 4.44 | 1.80 |
| 0 | -1 | 17 | 66.7 | 32.6 |
| 0 | -2 | 18 | 13.9 | 16.2 |
| 0 | -3 | 19 | 6.67 | 3.84 |
| -1 | -2 | 20 | 25.0 | 10.0 |
| -1 | -3 | 21 | 5.56 | 6.84 |
| | | | $\approx 310$ | $\approx 177$ |

$X_1$ = .881 resistive
$X_3$ = .012 conductive
$X_4$ = .012 conductive

Total error entering with negative sign $\approx 177$.

Pole-Dipole Measurements.
Pole Source at k
Dipole Receiver at (j, j-1).

$\beta_9$ and $\beta_{22}$ assume positive values, and errors corresponding to all other $\beta_m$ are still negative. Thus

$$\epsilon_i' = G_{i1}X_1'' - \beta_i > 0 \qquad i = 9, 22$$

$$\leq 0 \quad i = 1 \ldots 8, 10, \ldots 21$$

Next we equate the effects of cubes corresponding to $X_2, X_5, X_6$ to the major $\epsilon_i'$ values. This has two advantages; time of the second run is greatly reduced, and possibility of errors entering as positive errors is increased. A final solution can be easily deduced. Thus the second step is

$$G_{i2}X_2 + G_{i5}X_5 + G_{i6}X_6 - \beta_i' = G_{i1}X_1'' - \beta_i$$

for i's chosen from (1, ... , 22)

In the actual solution for the resistive cube, $X_1'' = .2$ was taken. It was noted that total error changed from $\approx 177$ to $\approx 70$. The second step was run and results are shown in Table III. The error column has been labelled as $\epsilon_i$ since now both positive and negative errors appear in the solution.

B. In this approach we choose an $X_1''$ such that when $X_1 \longrightarrow X_1'$, approximately a half of the $\epsilon_i'$ assume positive values. We incorporate this change of sign in the original equations. Thus finally we have a set of equations in which approximately half of error variables are positive and other half negative, and we re-run the whole program. The error will be greatly reduced. Residual error can further be reduced by fitting more cubes. This approach is used in the analysis of the conductive cube data.

TABLE III

Final Solution, Resistive Cube

| k | j | m | $b_m$ | $y_m$ |
|---|---|---|---|---|
| 4 | 2 | 1 | 5.56 | 0.71 |
| 4 | 1 | 2 | 3.33 | 0.10 |
| 4 | 0 | 3 | 7.50 | -1.73 |
| 4 | -1 | 4 | 0.555 | 0.08 |
| 3 | 1 | 5 | 2.78 | -1.51 |
| 3 | 0 | 6 | 14.4 | -0.91 |
| 3 | -1 | 7 | 2.22 | 0.62 |
| 3 | -2 | 8 | 1.39 | 0.20 |
| 2 | 1 | 9 | 8.33 | -0.52 |
| 2 | 0 | 10 | 27.8 | -4.08 |
| 2 | -1 | 11 | 7.22 | 1.35 |
| 2 | -2 | 12 | 0.555 | 0.320 |
| 2 | -3 | 13 | 1.11 | -0.17 |
| 1 | 0 | 14 | 75.0 | -2.86 |
| 1 | -1 | 15 | 25.0 | -1.37 |
| 1 | -2 | 16 | 4.44 | 1.00 |
| 0 | -1 | 17 | 66.7 | 0.00 |
| 0 | -2 | 18 | 13.9 | -2.25 |
| 0 | -3 | 19 | 6.67 | -0.29 |
| -1 | -2 | 20 | 25.0 | -0.35 |
| -1 | -3 | 21 | 5.56 | -2.11 |
| -2 | -3 | 22 | 8.33 | 0.00 |
| | | | $\approx 310$ | $\approx 20$ |

$X_1 = .681$ resistive $X_4 = .012$ conductive

$X_2 = .295$ conductive $X_5 = .145$ conductive

$X_3 = .012$ conductive $X_6 = .145$ conductive

Total absolute error entering with both positive and negative sign $\approx 20$.

Pole-Dipole Measurements - Pole Source at k
- Dipole Receiver at (j, j-1).

Conductive Cube

Following are the results of the analysis of the data, when the tank has a conductive (graphite) cube.

The data were analysed in two steps.

Step I.

Using $X_1$, $X_3$, $X_4$ the following L.P. problem was solved.

$$[G_{mi}][X_i] - [\beta_m] = [b_m]$$

Step II.

Approach B was used.

Results of Step I follow on the next page, Table IV. In Step I $X_1 = .924$. It was noted that when we put

$$X_1 = X_1' + X_2'' = .724 + .2 \qquad \text{i.e. } X_2'' = .2,$$

the errors

$$\beta_s ; \qquad s = 1, 2, 4, 7, 8, 9, 15, 18, 22$$

remain negative, while errors corresponding to

$$\beta_q ; \qquad q = 3, 5, 6, 10, 11, 12, 13, 14, 16, 17, 19, 20, 21$$

change sign and become positive. This change was incorporated and the program re-run. The error was reduced from $\approx 89$ to $\approx 39$. Thus with only cubes $X_1$, $X_3$, $X_4$ the error is less than 9%. We still have $X_2$, $X_5$, $X_6$, which we could use to bring the error down. This was, however, not done in the

TABLE IV

Partial Solution, Conductive Cube

| k | j | m | $b_m$ | $\beta_m$ |
|---|---|---|---|---|
| 4 | 2 | 1 | 2.78 | 3.47 |
| 4 | 1 | 2 | 7.78 | 2.05 |
| 4 | 0 | 3 | 14.2 | 1.89 |
| 4 | -1 | 4 | 0.866 | 0.266 |
| 3 | 1 | 5 | 13.9 | 0.5 |
| 3 | 0 | 6 | 25.0 | 3.1 |
| 3 | -1 | 7 | 2.50 | 1.25 |
| 3 | -2 | 8 | 1.67 | 0.53 |
| 2 | 1 | 9 | 11.1 | 8.2 |
| 2 | 0 | 10 | 50.0 | 7.3 |
| 2 | -1 | 11 | 8.67 | 0.00 |
| 2 | -2 | 12 | 1.67 | 0.28 |
| 2 | -3 | 13 | 1.94 | 0.25 |
| 1 | 0 | 14 | 97.2 | 14.2 |
| 1 | -1 | 15 | 36.1 | 11.3 |
| 1 | -2 | 16 | 4.16 | 0.16 |
| 0 | -1 | 17 | 77.8 | 19.8 |
| 0 | -2 | 18 | 29.2 | 7.9 |
| 0 | -3 | 19 | 9.44 | 1.03 |
| -1 | -2 | 20 | 22.2 | 3.0 |
| -1 | -3 | 21 | 16.7 | 2.4 |
| -2 | -3 | 22 | 2.78 | 0.00 |
| | | | $\approx 440$ | $\approx 89$ |

$X_1$ = .924 conductive
$X_3$ = .032 resistive
$X_4$ = .032 resistive

Total error entering with negative sign $\approx 89$

Pole-Dipole Measurements - Pole Source at k
- Dipole Receiver at (j, j-1).

present work. In the final tabulated results, the error column is labelled as $\epsilon_m$ , showing both positive and negative errors entered in the solution (Table V).

Approach A could be very useful, if it was required to incorporate information from other sources into the analysis of field measurements. A feature of this approach is that the time required for computing the second run is much less than the time required for computing the initial run. In case of small problems this can come in very handy.

Approach B is to be identified with reducing the absolute numerical value of $[C_m][\alpha_m - \beta_m]$ . It more or less gives symmetrical fitting of the data. A feature of this approach is that we do not subjectively change the value of $X_1$, but this results in its becoming more time consuming than approach A. This may not be justified in most cases.

The computation time using the L.P. approach is just about the same as the time required when the method of least squares is used. For comparison, a least squares solution has been done on the conductive cube using the same three contributors that were used in the L.P. solution. This is shown in Appendix II.

TABLE V

Final Solution, Conductive Cube*

| k | j | m | $b_m$ | $\epsilon_m$ |
|---|---|---|---|---|
| 4 | 2 | 1 | 2.78 | -2.20 |
| 4 | 1 | 2 | 7.78 | 0.00 |
| 4 | 0 | 3 | 14.2 | 1.4 |
| 4 | -1 | 4 | 0.866 | -0.023 |
| 3 | 1 | 5 | 13.9 | 2.5 |
| 3 | 0 | 6 | 25.0 | 2.7 |
| 3 | -1 | 7 | 2.50 | -0.43 |
| 3 | -2 | 8 | 1.67 | -0.08 |
| 2 | 1 | 9 | 11.1 | -4.1 |
| 2 | 0 | 10 | 50.0 | 4.5 |
| 2 | -1 | 11 | 8.67 | 1.73 |
| 2 | -2 | 12 | 1.67 | 0.12 |
| 2 | -3 | 13 | 1.94 | 0.20 |
| 1 | 0 | 14 | 97.2 | 8.4 |
| 1 | -1 | 15 | 36.1 | -1.5 |
| 1 | -2 | 16 | 4.17 | 0.70 |
| 0 | -1 | 17 | 77.8 | 0.00 |
| 0 | -2 | 18 | 29.2 | -0.2 |
| 0 | -3 | 19 | 9.44 | 1.09 |
| -1 | -2 | 20 | 22.2 | 2.0 |
| -1 | -3 | 21 | 16.7 | 3.5 |
| -2 | -3 | 22 | 2.78 | -1.41 |

*Compare with least squares solution in Appendix II. $\approx 440$ $\approx 39$

$X_1 = .734$ conductive

$X_3 = .025$ resistive

$X_4 = .025$ resistive

Total absolute value of error entering with both positive and negative sign $\approx 39$

Pole-Dipole Measurements - Pole Source at k
- Dipole Receiver at (j, j-1).

## VII. Discussion of Actual Computations

The geometric factors $G_{ijk}$ are surface integrals which appear in the linear approximation. These were calculated numerically [7], and are accurate to about 1%. The factors take on both positive and negative signs. As a result of the approximation involved in their determination there are regions (j,k) where their zeros do not coincide with the zeros of the measured data. The computed values of $G_{ijk}$ refer to pole-pole configuration of the source and receiver. However, pole-dipole data were analysed and the $G_{ijk}$ were used to calculate the coefficients for the pole-dipole configuration. Accordingly probability of a geometrical coefficient having a wrong sign was increased.

Using only $X_1$ in the analysis it was seen that the error could not be reduced to the order of 5% - 10%, hence the basic hypothesis of dividing the region into cubes had to be evoked in greater degree. Now another interesting observation was made. When the program was run with all six cubes available, only one or two showed up in the solution. For effective reduction of error it is necessary that the cubes other than $X_1$ should have finite conductivities, slightly differing from the background. If they show up as having the

same resistivity as the background, the corresponding X will show up as zero in the solution, and the error will not be decreased. There seems to be a reason to explain this. As far as the mathematical structure of the system of equations is concerned, a wrong sign of the coefficient of X in any equation will affect the whole solution, and will either give X a very wrong value or will simply make it zero. Obviously the nonappearance of certain X's in the solution is due to some anomaly in signs of the coefficients of X. That some anomalies are present is quite obvious and in some measure may be attributed to uncertainty in sign of coefficients $G_{ijk}$. However, it was also noted that anomaly in sign usually showed up when the spacing between the source and receiver was large, i.e. when the larger volumes are being explored. It appears plausible that perhaps some surface and boundary effects come into play. Thus the discrepancy in sign may be tied in with the size of the tank and perhaps in a bigger tank these would not appear at all. On the basis of the above arguments, signs of four to six coefficients were arbitrarily corrected in the complete analysis of a set of data. It was also noted that if conductivities $X_3$ and $X_4$ were forced to be equal, and if $X_5$ and $X_6$ were also forced to be equal, the number of sign discrepancies was reduced.

## VIII. Discussion of Results

The method is quite successful in semiquantitatively analysing the problem. The semiquantitative nature of interpretation arises, not from the method of fitting used, but from the nature of contrast factor $X_i$, and the linear approximation itself. We notice that

$$C X_i = C \frac{\sigma_i - \sigma_0}{\sigma_i + 2\sigma_0}, \quad C = 3.6$$

is not a very sensitive parameter as is remarked in [7] The parameter $X_i$ approaches 1 as conductivity increases, and approaches $-\frac{1}{2}$ as resistivity increases. In case of conductive cube a value of .73 is noted, while in case of resistive cube a value of -.68 is noted, for this parameter. The absolute value $.68 > .5$ thus does not fall within the allowable range of $X_i$. However, it must be remembered that the parameter $C \frac{\sigma_i - \sigma_0}{\sigma_i + 2\sigma_0}$ is entirely empirical and some deviations from it are expected as may be clearly seen from the graph used to determine $C = 3.6$ [7].

A tremendous improvement over least square fitting is not noted. The work of Ness [4] shows that the least square method of fitting at best gives results in the range of 5% - 10%. The results achieved by L.P. methods are clearly

in this range*. Hence the latter methods are at least as good as least squares and due to their flexibility may even claim an advantage over the method of least squares. However, for large problems L.P. methods tend to consume more and more time.

The aim of achieving something better than 5% error in fitting is, from the point of view of the actual errors involved, not justified. First of all an error of about 1% exists in the computed coefficients $G_{ijk}$. Then there are experimental errors which are of the order of (2-3)% at least. Finally an unknown amount of error is introduced in the linear approximation and in the empirical parameter $C\ \frac{\sigma_i - \sigma_0}{\sigma_i + 2\sigma_0}$. These errors on the whole are sufficient to limit the solution to a 5% error fit.

However, the purpose of this work was to see whether or not L.P. could be used. That it can be used is demonstrated.

---

*For example, compare result of Table V with that in Appendix II, Table VI.

## IX. Conclusions

The results obtained show that the L.P. methods can be used for the interpretation of resistivity data. From the point of view of the computing time and accuracy involved, these methods appear to be as good as the method of least squares, at least on these two examples. The L.P. methods are quite flexible and provide means of subjective control over the solution. In cases where the latter is necessary, these methods are more powerful than the method of least squares, and can be gainfully employed. At the present stage, however, the choice between the two methods would perhaps be dictated by the amount of data to be interpreted, by the type of information desired, and by the amount of time the analyst has available.

## X. Suggestions for Further Work

The problem of resistivity interpretation can be formulated as a problem in Theory of Games. It may be mentioned here that Theory of Games and Linear Programming are closely related to each other.

We may consider, in principle, that a list has been drawn up of every possible state of the matrix of cubes into which the subsurface region has been subdivided. Each cube can exist in one of three states (assumed equally likely) i.e. it can be more conductive than background ( $\sigma_i > \sigma_0$ ), less conductive than background ( $\sigma_i < \sigma_0$ ) or same as background ( $\sigma_i = \sigma_0$ ). Thus correspondingly the matrix of cubes can exist in $3^n$ states. In the terminology of Theory of Games, each state of the matrix of cubes is termed as a pure strategy of nature. The geophysicist makes m measurements on the surface of the earth. He then attempts to predict the state of the matrix. His predictions involve the state of the matrix of cubes, and also the experimental error, which can be either positive, negative, or zero. Thus we may regard the geophysicist as having $3^{n+m}$ strategies. The aim is to make predictions in an optimal way. Games of this type are discussed in "Introduction to the Theory of Games" by

McKinsey (McGraw-Hill). However, it is obvious that the above formulation will have to be suitably modified before a practical solution can be devised.

Another interesting suggestion is due to Vozoff. When the graph used to determine $C = 3.6$ in [7] is closely examined, it is clear that the parameter $\frac{\sigma_i - \sigma_0}{\sigma_i + 2\sigma_0}$ is either $-\frac{1}{2}$ or $1$ for most parts of the range of the ratio $\frac{\sigma_i}{\sigma_0}$. It shows variation only in a very short range when $\frac{\sigma_i}{\sigma_0} \approx 1$. Thus if empirically it is hypothesized that $\frac{\sigma_i - \sigma_0}{\sigma_i + 2\sigma_0}$ is either one value of the set of numbers $\{-\frac{1}{2}, 0, 1\}$ then under suitable assumptions, a formulation in terms of game theory may be possible. Statistical games and testing of hypotheses are discussed in "Theory of Games and Statistical Decisions" by Blackwell and Girshick (John Wiley). In many cases resistivity methods are used to decide whether a subsurface region is resistive or conductive relative to background without reference to actual magnitude of $\sigma_i$ In such cases the above formulation would be justified.

A realistic linear programming model requires accurate and reliable numerical values of the coefficients $G_{ijk}$. Hence it would be useful to study the behaviour of solutions when these coefficients are allowed to vary.

## Appendix I

### The Algebraic Expression for the Geometrical Coefficients $G_{ijk}$ and Some Numerical Values Obtained Therefrom

In the resistivity approximation coefficients $G_{ijk}$ for pole-pole configuration are given by

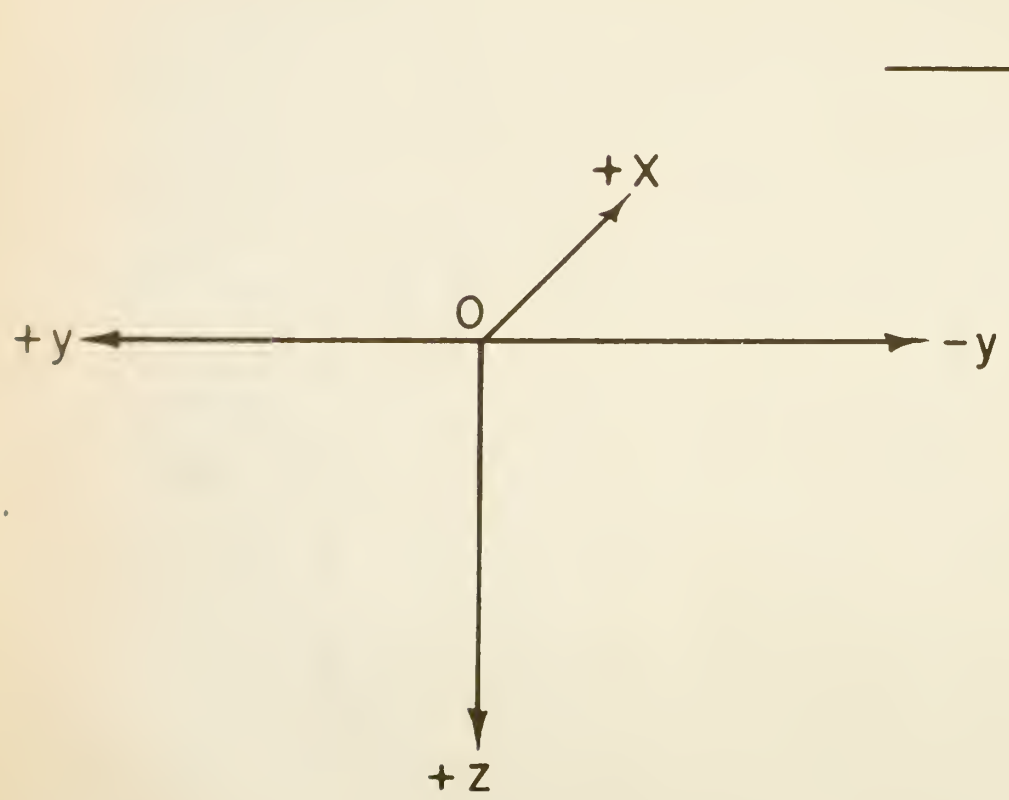
$$G_{ijk} = \left\{ \iint_{S_1} \frac{\hat{i}\cdot\vec{r}_{RQ}}{r_{RP}\, r_{RQ}^3}\, dydz - \iint_{S_2} \frac{\hat{i}\cdot\vec{r}_{RQ}}{r_{RP}\, r_{RQ}^3}\, dydz \right\}$$

$$+ \left\{ \iint_{S_3} \frac{\hat{j}\cdot\vec{r}_{RQ}}{r_{RP}\, r_{RQ}^3}\, dzdx - \iint_{S_4} \frac{\hat{j}\cdot\vec{r}_{RQ}}{r_{RP}\, r_{RQ}^3}\, dzdx \right\}$$

$$+ \left\{ \iint_{S_6} \frac{\hat{k}\cdot\vec{r}_{RQ}}{r_{RP}\, r_{RQ}^3}\, dxdy - \iint_{S_5} \frac{\hat{k}\cdot\vec{r}_{RQ}}{r_{RP}\, r_{RQ}^3}\, dxdy \right\}$$

(See Fig. 4)

GEOMETRICAL PICTURE FOR A SINGLE CUBE



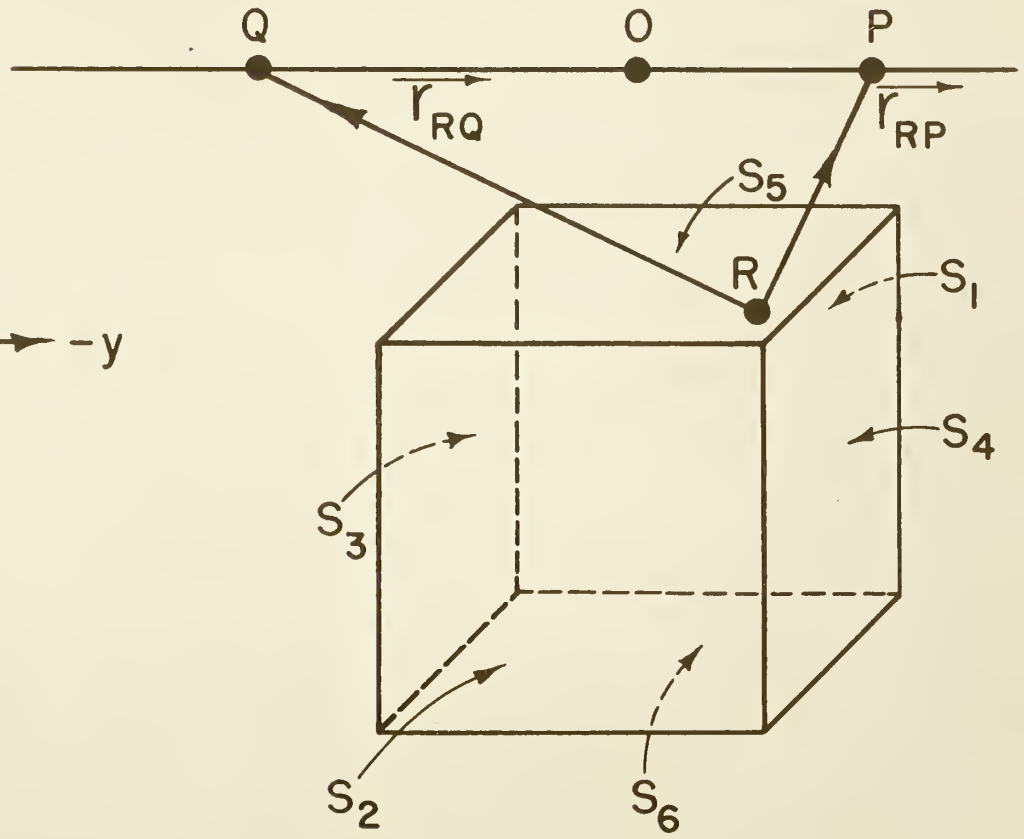




FIG. 4

The coefficients $G_{ijk}$ were computed numerically on an electronic computer [7]. From these, the geometrical coefficients corresponding to pole-dipole configuration were computed for the cases when the centre of the disturbing cube was 1 unit and 2 units below the surface of measurement. These numerical values are shown graphically on the following page. (See Figs. 5 and 6.)

# GRAPHICAL PRESENTATION OF DATA

## POLE – DIPOLE GEOMETRICAL COEFFICIENTS

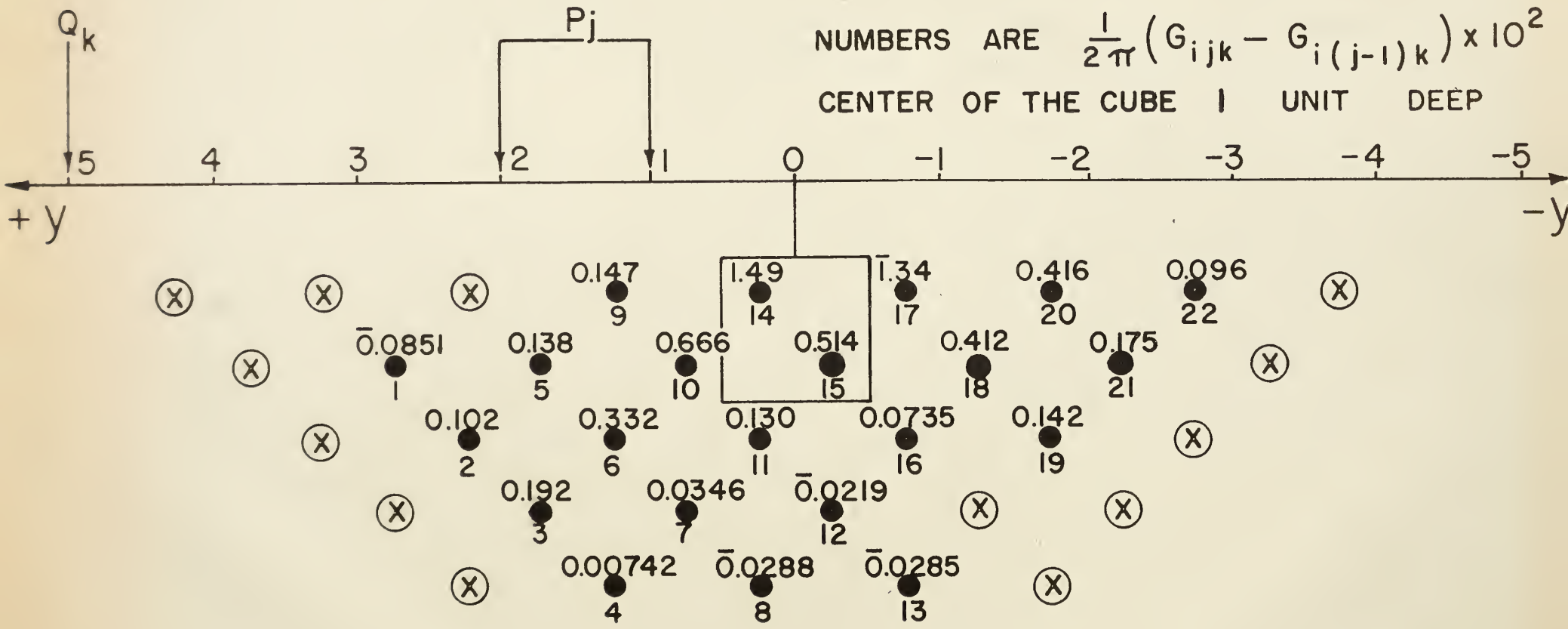



FIG. 5

# GRAPHICAL PRESENTATION OF DATA

## POLE-DIPOLE GEOMETRICAL COEFFICIENTS

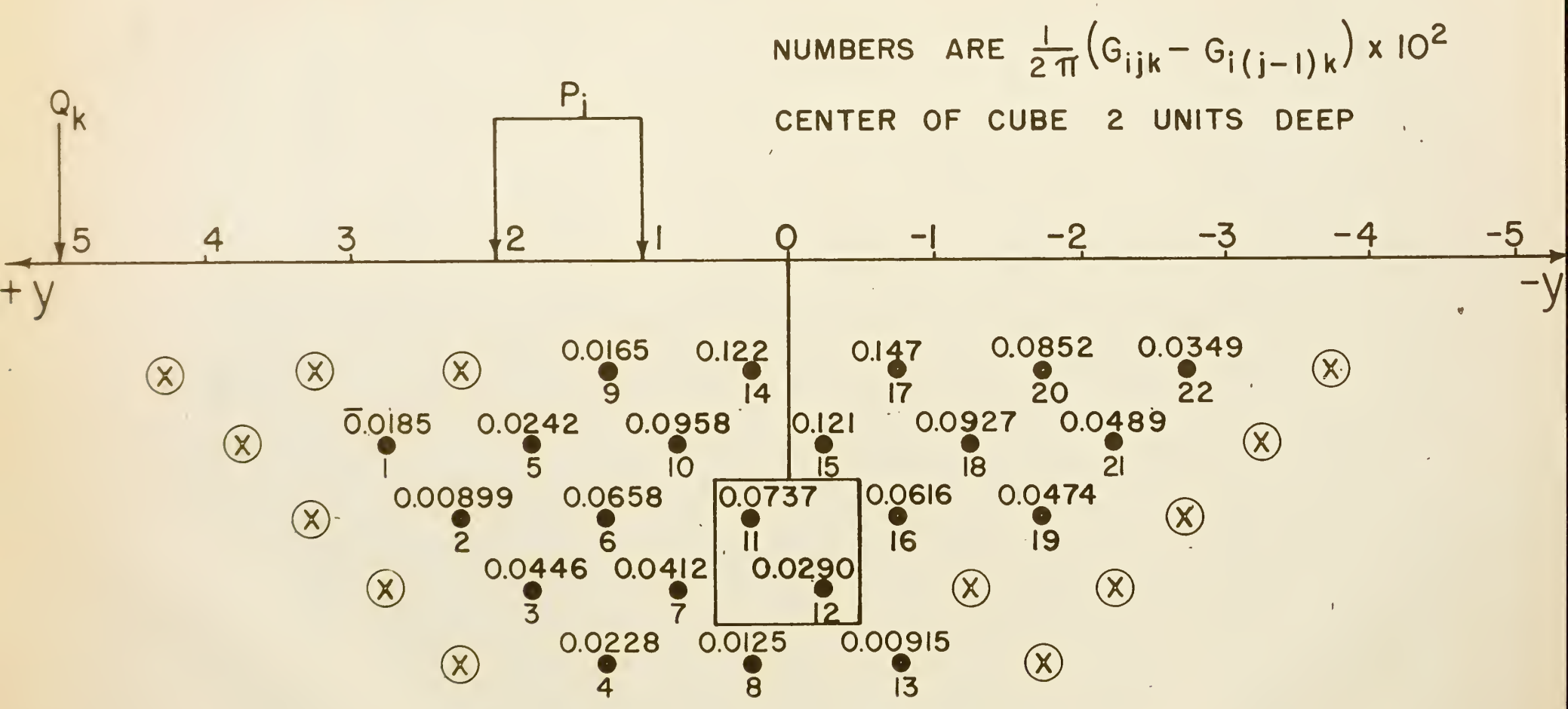



FIG. 6

## Appendix II

### Least Squares Solution, Conductive Cube

Model data due to a single conductive cube (See Table V) were analysed by the method of least squares. The results of this analysis are shown in Table VI. The results obtained by the two methods (Table V, Table VI) are in the same range of accuracy.

TABLE VI

Least Squares Solution, Conductive Cube

| k | j | m | $b_m$ | $\epsilon_m$ |
|---|---|---|---|---|
| 4 | 2 | 1 | 2.78 | -2.41 |
| 4 | 1 | 2 | 7.78 | 0.18 |
| 4 | 0 | 3 | 14.2 | -0.4 |
| 4 | -1 | 4 | 0.867 | 0.060 |
| 3 | 1 | 5 | 13.9 | 4.6 |
| 3 | 0 | 6 | 25.0 | 2.7 |
| 3 | -1 | 7 | 2.50 | -0.34 |
| 3 | -2 | 8 | 1.67 | -0.13 |
| 2 | 1 | 9 | 11.1 | -3.0 |
| 2 | 0 | 10 | 50.0 | 4.5 |
| 2 | -1 | 11 | 8.67 | 1.22 |
| 2 | -2 | 12 | 1.67 | -0.12 |
| 2 | -3 | 13 | 1.94 | 0.01 |
| 1 | 0 | 14 | 97.2 | 5.6 |
| 1 | -1 | 15 | 36.1 | -1.0 |
| 1 | -2 | 16 | 4.17 | 0.37 |
| 0 | -1 | 17 | 77.8 | -3.0 |
| 0 | -2 | 18 | 29.2 | 1.10 |
| 0 | -3 | 19 | 9.44 | 0.81 |
| -1 | -2 | 20 | 22.2 | 0.20 |
| -1 | -3 | 21 | 16.7 | 3.74 |
| -2 | -3 | 22 | 2.78 | -1.92 |
| | | | $\approx 440$ | $\approx 36$ |

$X_1 = 0.746$ conductive

$X_3 = 0.020$ resistive

$X_4 = 0.020$ resistive

Total absolute value of error entering with both positive and negative sign $\approx 36$

Pole-Dipole Measurements - Pole Source at k
- Dipole Receiver at (j, j-1).

## References


1. Burnside, W. (1959) "Theory of Probability," Dover Publications Inc., New York.

2. Fisher, R. A. (1950) "Contributions to Mathematical Statistics," John Wiley & Sons, Inc., New York.

3. Gass, S. I. (1958). "Linear Programming," McGraw-Hill Book Company Inc., New York.

4. Ness, N. (1959) "Resistivity interpretation in geophysical prospecting," Ph.D. Thesis: Department of Geology and Geophysics, Massachusetts Institute of Technology.

5. Whittaker, E. and Robinson, G. (1954) "Calculus of Observation," Blackie & Sons Ltd., London.

6. Vozoff, K. (1956) "Quantitative analysis of earth resistivity data," Ph.D. Thesis: Massachusetts Institute of Technology.

7. Vozoff, K. (1960) "Numerical resistivity interpretation: general inhomogeneity," Geophysics, XXV, 6.

8. Saaty, T. L. (1959) "Mathematical Methods of Operations Research," McGraw-Hill Book Co. Ltd., New York.

9. Vzjda, S. (1954) "Theory of Games and Linear Programming," Methuen & Co. Ltd., London.


## Additional Bibliography


Blackwell, D. and Girshick, M.A. (1954) "Theory of Games and Statistical Decisions," John Wiley & Sons Inc., New York.

Hildebrand, F. B. (1956) "Introduction to Numerical Analysis," McGraw Hill Book Co. Ltd., New York.

Hildebrand, F. B. (1952) "Methods of Applied Mathematics," Prentice Hall Inc., New York.

McKinsey, J. C. C. (1952) "Introduction to Theory of Games," McGraw-Hill Book Co. Ltd., New York.

Marshall, D. J. and Madden, T. R. (1959) "Induced polarisation, a study of its causes," Geophysics, XXIV, 5.

Topping, J. (1958) "Errors of Observation and their Treatment," Chapman and Hall Ltd., London.